

# Quality improvement of adaptive optics retinal images using conditional adversarial networks

WANYUE LI,<sup>1,2,3,5</sup>  GUANGXING LIU,<sup>3,5</sup> YI HE,<sup>2,3,6</sup>  JING WANG,<sup>1,2,3</sup> WEN KONG,<sup>1,2,3</sup> AND GUOHUA SHI<sup>1,2,3,4,7</sup> 

<sup>1</sup>University of Science and Technology of China, Hefei, 230041, China

<sup>2</sup>Jiangsu Key Laboratory of Medical Optics, Suzhou, 215163, China

<sup>3</sup>Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, 215163, China

<sup>4</sup>Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, 200031, China

<sup>5</sup>These authors contributed to the work equally and should be considered co-first authors

<sup>6</sup>heyi@sibet.ac.cn

<sup>7</sup>ghshi\_lab@126.com

**Abstract:** The adaptive optics (AO) technique is widely used to compensate for ocular aberrations and improve imaging resolution. However, when affected by intraocular scatter, speckle noise, and other factors, the quality of the retinal image will be degraded. To effectively improve the image quality without increasing the imaging system's complexity, the post-processing method of image deblurring is adopted. In this study, we proposed a conditional adversarial network-based method for directly learning an end-to-end mapping between blurry and restored AO retinal images. The proposed model was validated on synthetically generated AO retinal images and real retinal images. The restoration results of synthetic images were evaluated with the metrics of peak signal-to-noise ratio (PSNR), structural similarity (SSIM), perceptual distance, and error rate of cone counting. Moreover, the blind image quality index (BIQI) was used as the no-reference image quality assessment (NR-IQA) algorithm to evaluate the restoration results on real AO retinal images. The experimental results indicate that the images restored by the proposed method have sharper quality and higher signal-to-noise ratio (SNR) when compared with other state-of-the-art methods, which has great practical significance for clinical research and analysis.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Retinal imaging is one of the most useful modalities of clinical research and diagnosis in ophthalmology. However, direct observation of the retina is inevitably affected by ocular aberrations and the imaging resolution is severely limited. To compensate for the ocular aberrations, the adaptive optics (AO) technique was introduced to get nearly diffraction-limited resolution [1,2]. Since then, AO has helped us achieve microscopic imaging of the living human retina at the single-cell level and has been successfully integrated into the confocal scanning laser ophthalmoscope (SLO) to improve its imaging resolution of the human retina [3–5]. However, the ocular aberrations compensation is not perfect, and it still leaves some residual aberration. Furthermore, affected by intraocular scatter, uncontrolled physiological vibration of eye, speckle noise, and other factors, the quality of adaptive optics confocal scanning laser ophthalmoscope (AOSLO) images is further degraded and the photoreceptor cells become obscured. Therefore, an appropriate image post-processing method is indispensable to enhance the retinal images, which can help better detect the photoreceptor cells and assist clinicians in the examination of the human retina.

Image deblurring is an effective and widely used post-processing method for image quality improvement. Algorithms that tackle image deblurring tasks can be mainly classified as either deconvolution-based or learning-based methods. Deconvolution-based methods can be further divided into non-blind and blind deconvolution methods. Early work mostly focused on non-blind deconvolution methods, making an assumption that the point-spread functions (PSFs) are known. For AO retinal image deblurring, Arines et al. [6] measured the PSFs using a wavefront sensor to compensate for ocular residual aberrations left by AO. Although this method can be used in nearly real time, its performance heavily depends on the accuracy of PSFs' measurement. To make up for the limitations of the non-blind deconvolution method, a more generalized, blind deconvolution method has been proposed. This type of method allows for recovery of both PSFs and target images, and has been widely used in AO retinal image deblurring tasks [7–9]. In [7], the authors introduced blind deconvolution methods either in combination or as a substitute for AO in a scanning laser ophthalmoscope. Christou et al. [8] used a blind deconvolution algorithm based on multi-frame images to recover the AO retinal images and the PSFs. In [9], the authors used the incremental Wiener filter as a blind deconvolution technique to restore the AOSLO retinal images and corresponding PSFs. However, the disadvantage of the blind deconvolution method is that it often gets trapped in local minima, which makes it hard to find a unique solution, especially when there is only a single-blurred image to be restored [10,11].

Recently, the deep learning-based method has been introduced to enhance AO retinal images [10,12]. In [12], the authors employed the random forest to learn the mapping of retinal images onto the space of blur kernels expressed in terms of Zernike coefficients. For an input image, the optimal vector of the Zernike coefficient is predicted by the trained random forest, and then the corresponding PSF is reconstructed. However, this method cannot be referred to as a fully deep-learning-based method; this method only uses deep learning to predict PSFs and still need non-blind deconvolution to restore the blurred images. Besides, this method was designed for the retinal images captured from a commercially available flood-illuminated AO instrument (rtx1, Imagine Eyes, Orsay, France); for other AO systems, this method may fail. Furthermore, this method neglects the noise term and assumes that images are corrupted only by convolution blur kernels. X. Fei et al. [10] proposed a fully learning-based method, and a five-layers convolutional neural network (CNN) was used to learn an end-to-end mapping between blurred and restored retinal images. For image restoration problems, image details are very important. The method proposed in [10] only uses five convolutional layers and cannot extract features rich and varied enough, which may affect the quality of restored images. Furthermore, lots of studies [13–16] have verified that L2 loss usually leads to blurry artifacts on generated images, this method uses L2 loss as cost function will blur the image details up to a point. Moreover, this model was only trained on space-invariant PSFs with Gaussian noise, which may affect the generalization ability of the models to some extent.

In this study, we proposed a generative adversarial network (GAN) [17]-based method for AO retinal image deblurring. GAN-based methods are new to the AO retinal image deblurring task, but have already shown promise in natural scene image deblurring [18,19]. In this approach, the conditional GAN [20] was used to directly learn the mapping relationship between blurry and restored retinal images. The proposed network is composed of a generator network and a discriminator network. The generator network is a 16-layer symmetric deep CNN that can extract image features as rich and varied as possible. The discriminator network is inspired by the PatchGAN [21,22] model, which penalizes the image structure at the scale of patches instead of the whole images, and can help reconstruct image details and improve the training efficiency. Moreover, we introduced pixel-space loss (L1 loss) and feature-space loss (VGG loss) in the training of the proposed model, which can ensure the restored image's high signal-to-noise ratio (SNR) and better visual effect. The proposed network was validated on synthetic and real AO retinal images, and the assessment results demonstrate the effectiveness of the proposed method

in image quality improvement. In summary, compared with previous methods [10,18,23], the proposed method is well generalized and can effectively improve the image quality, especially with large residual wavefront aberrations and noise. The rest of this study is structured as follows: in section 2, we discuss the datasets and the proposed method. The results of our method, and its quantitative and qualitative evaluation, are described in section 3. In section 4, discussions of this study are given and conclusions of our work are drawn.

## 2. Method

### 2.1. Datasets

The dataset used for training is very important for learning-based methods. To make the network work well, large-scale training data are needed, however, a large amount of well-corrected and uncorrected AO retinal image pairs are difficult to obtain, and the images captured with AO are still affected by noises and residual wavefront aberrations. Therefore, like previous learning-based methods [10,12], our model was also trained on synthetically generated retinal images.

In [10] and [12], the authors proposed approaches for synthetic AO retinal images generation. Both approaches are comprised of four key steps: (1) create a set of ideal retinal images using the algorithm proposed in [24], (2) generate a set of PSFs to simulate the residual optical aberrations of the eye, (3) convolve the ideal retinal images with PSFs to generate synthetic images, and (4) add Gaussian noise to the generated images.

The PSFs to simulate the residual optical aberrations of the eye are described as follows:

$$\text{PSF} = \left\| \text{FFT} \left\{ P(x, y) \exp \left( -i \frac{2\pi}{\lambda} \varphi(x, y) \right) \right\} \right\|^2 \quad (1)$$

$$P(x, y) = \begin{cases} 1 & x^2 + y^2 \leq r^2 \\ 0 & \text{other} \end{cases} \quad (2)$$

$$\varphi(x, y) = \sum_i a_i Z_i(x, y), \quad (3)$$

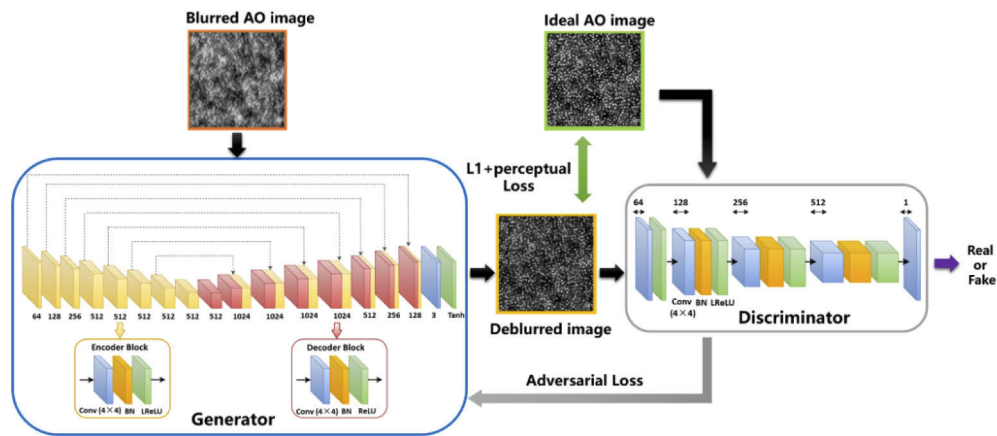
where  $P$  is the pupil function,  $r$  is the pupil diameter,  $\lambda$  is the wavelength of the imaging beam,  $\varphi(x, y)$  is the wavefront phase error,  $a_i$  and  $Z_i$  are the Zernike coefficient and Zernike polynomial, respectively, and  $i$  is set from 3 to 35.

In this work, we also use the approach described above. The values of the Zernike coefficients  $a_i$  were selected from a statistical model of wavefront aberrations in healthy eyes reported in [25]. For a retinal imaging AO system, the quantity of residual aberrations after correction is usually small [26]. To simulate the correction results of AO, the low-order aberrations were set to zero and the values of high-order Zernike coefficients were scaled by 0.1. For most research and commercial AO instruments [3,27–30], the wavelength of them usually ranges from 650 nm to 850 nm, and the pupil diameter is usually chosen from 6 mm to 8 mm. To ensure that the trained network can be widely used in different AO systems, the wavelength  $\lambda$  in this study was assigned from 650 nm to 850 nm with a step of 10 nm, and the pupil diameter  $r$  was between 6 mm and 8 mm. By doing so, 150 ideal retinal images with the field of view of 1.3 degrees and different eccentricities (range from 0.3 mm to 1.5 mm from the foveal center) were created, and 4,000 PSFs were generated. To train and evaluate the model, 25 ideal retinal images with different eccentricities were randomly selected, and each of them was convolved with 4,000 PSFs, half of which were used for validation to monitor overfitting during training, and the other half were used for testing. The remaining 125 ideal retinal images were also convolved with each of the 4,000 PSFs, and used for model training. Thus, the training, validation, and testing dataset was composed of 500,000, 50,000, and 50,000 image pairs, respectively. In addition, Gaussian noise

with standard deviation uniformly sampled from  $[0, 0.05]$  was added to half of the synthetic images in each dataset.

## 2.2. Conditional adversarial network for AO retinal image enhancement

To directly learn a mapping from an input blurry AO retinal image to an ideal retinal image, we proposed a conditional GAN-based deep network. As described in Fig. 1, this proposed network is mainly composed of two important parts (generator G and discriminator D). The generator network is a symmetric deep CNN network with skip connections whose structure is similar to that of U-net, as shown in the left part of Fig. 1. The primary goal of the generator is to generate a deblurred image from an input blurry AO retinal image. The discriminator, as shown in the right part of Fig. 1, serves to distinguish the generated image from the corresponding real image and can also be viewed as a guide for the training of generator.



**Fig. 1.** Proposed model architecture. Our model is composed of a U-net based generator and a PatchGAN-based discriminator. (conv: convolutional layer; BN: batch normalization layer; ReLU: rectified linear unit; LReLU: leaky rectified linear unit; Tanh: TanHyperbolic function)

In order to learn a good generator G to fool the learned discriminator D and to make the discriminator good enough to distinguish the generated image from the real image, we introduce the Wasserstein distance [31] as an indicator of the training process. The Wasserstein distance has the desirable property of being continuous and differentiable almost everywhere under mild assumptions, and is informally defined as the minimum cost of transforming the distribution of generated images to the distribution of target images. Therefore, using the Wasserstein distance to measure the difference between the generated image and the ideal image can prevent gradient vanishing and enable us to obtain better restoration results [32].

The main goal of the proposed model is to learn a mapping function from an input blurry AO image  $I_B$  and a random noise vector  $z$  to the real AO image  $I_R$  by solving the following min-max problem:

$$\min_G \max_D E_{I_B \sim p(I_B), z \sim p(z)} - D(I_B, G(I_B, z)) + E_{I_B \sim p(I_B), I_R \sim p(I_R)} D(I_B, I_R) \quad (4)$$

The proposed generator G with a symmetric structure is illustrated in the left part of Fig. 1. The core of the generator is a series of encoder and decoder blocks. Each encoder and decoder block is composed of a convolutional layer with a kernel size of  $4 \times 4$  and a stride of  $2 \times 2$ . Each convolutional layer is followed by a batch normalization (BN) layer and an activation layer. The difference between the encoder and decoder block is that the activation function of the decoder

block is a rectified linear unit (ReLU), and that of the encoder block is a leaky ReLU (LReLU). All of the convolutional layers in the encoder are downsampled by a factor of 2, whereas those in the decoder they are upsampled by a factor of 2. After the last decoder block, a convolution layer is applied to map the number of output channels; this is followed by a Tanh function. The U-net architecture is identical except with skip connections between each block  $i$  in the encoder and block  $n-i$  in the decoder, where  $n$  is the total number of convolutional layers. The skip connections concatenate activations from layer  $i$  to layer  $n-i$ , and change the number of channels in the decoder.

The discriminator of the traditional GAN only distinguishes the generated images and real images based on their overall information, which will lead to serious loss of image information [21]. Inspired by the idea behind PatchGAN, we designed a discriminator as shown in the right part of Fig. 1. This network is composed of five convolutional layers, and the second to fourth convolutional layers are followed by a batch normalization layer and an LReLU activation function. The last convolutional layer is used to map to a 1-dimensional output. Using this network as a discriminator can help us recover image details and improve the training efficiency.

### 2.3. Loss function

To ensure that the generated images resemble the label images, we formulate the loss function as a combination of adversarial, pixel-space, and feature-space loss:

$$L = L_{GAN} + \alpha L_{pixel} + \beta L_{perceptual}, \quad (5)$$

where  $\alpha$  and  $\beta$  are the experimentally determined hyperparameters that control the effect of pixel-space and feature-space loss. Pursuing the balance between pixel-space and feature-space loss is not a trivial task. Giving over-weight to feature-space loss will result in the loss of image intensity information after delurring [18]; thus, pixel-space loss should have a higher weights than feature-space loss. In this task, we set  $\alpha=100$  and  $\beta=0.001$ , which were found to be the suitable parameters to achieve better performance.

**Adversarial loss:** Most of the studies related to conditional GANs use the vanilla GAN [33,34] objective as the loss function. In this study, we use WGAN [31] as the adversarial loss function since it is more stable and can generate higher quality results. The loss is calculated as follows:

$$L_{GAN} = \sum_{n=1}^N -D_{\theta_D}(G_{\theta_G}(I_B)). \quad (6)$$

**Pixel-space loss:** Common choices for the pixel-space loss function are L1 or L2 loss. Since L2 loss usually leads to blurry artifacts in generated images [13–16], we use L1 loss as the pixel-space loss to ensure the accurate restoration of pixel-level information. L1 loss is defined as

$$L_{pixel} = \frac{1}{WH} \sum_{x=1}^W \sum_{y=1}^H |(I_R)_{x,y} - (G_{\theta_G}(I_B))_{x,y}|, \quad (7)$$

where  $W$  and  $H$  are the width and height of the image, respectively.

**Feature-space loss:** Only comparing images in pixel-space is not enough to recover textural information. Perceptual loss is used as the feature-space loss, which is based on the difference of generated and target image CNN feature maps, and can help generate high perceptual quality images. It can be explained as

$$L_{perceptual} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I_R)_{x,y} - \phi_{i,j}(G_{\theta_G}(I_B))_{x,y})^2, \quad (8)$$



where  $\phi_{i,j}$  is the feature map obtained by the  $j$ th convolution before the  $i$ th maxpooling layer within the VGG19 network, pretrained on ImageNet [35], and  $W_{i,j}$  and  $H_{i,j}$  are the dimensions of  $\phi$ .

### 3. Experimental results

The Pytorch 0.4.1 library was employed to train and test the proposed method. All experiments were carried out in an Ubuntu 16.04+python3.6 environment. The program ran on a desktop computer with an Intel Xeon E5-2620 CPU with 32 GB RAM, and two NVIDIA GeForce GTX 1080Ti GPUs. We trained the proposed model for 100 epochs using the Adam [36] optimizer. The initial learning rate was set to 0.0002 for both the generator and the discriminator. After the first 50 epochs, the learning rate was linearly decayed to zero over the next 50 epochs. The proposed model was trained with a batch size equal to 1, which showed empirically better results on validation. It took nearly 16 hours to train the proposed model.

#### 3.1. Results evaluation for synthetic retinal images

We tested the proposed model on the synthetic retinal image dataset. This test was conducted on 50,000 AO retinal images of our test set, with the criteria of peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and perceptual distance [37]. The PSNR and SSIM are two commonly used metrics for quantitatively evaluating image restoration quality, and are described as follows:

$$MSE = \frac{\sum_{n=1}^N (x^n - y^n)}{N} \quad (9)$$

$$PSNR = 10 \times \lg \left( \frac{255^2}{MSE} \right) \quad (10)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (11)$$

where  $MSE$  is the mean square error;  $N$  is the size of image;  $x^n$  and  $y^n$  are the  $n$ th pixels of original image  $x$  and processed image  $y$ ;  $\mu_x, \mu_y$  are the averages of  $x, y$ ;  $\sigma_x^2, \sigma_y^2$  are the variance of  $x, y$ ; and  $\sigma_{xy}$  is the covariance of  $x$  and  $y$ .

Perceptual distance is used for measuring the perceptual quality of images, and partially coincides with the human perception. A smaller perceptual distance implies better perceptual quality and visual effects [38]. Given a reference image patch  $x$  and a distorted patch  $y$ , the perceptual distance can be calculated as follows:

$$d(x, y) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (\hat{x}_{hw}^l - \hat{y}_{hw}^l)\|_2^2, \quad (12)$$

where  $\hat{x}_{hw}^l$  and  $\hat{y}_{hw}^l$  are the feature stacks extracted from layer  $l$  of a specific network and unit-normalized in the channel dimension;  $w_l$  is the scale vector;  $H_l$  and  $W_l$  are the dimensions of image patches.

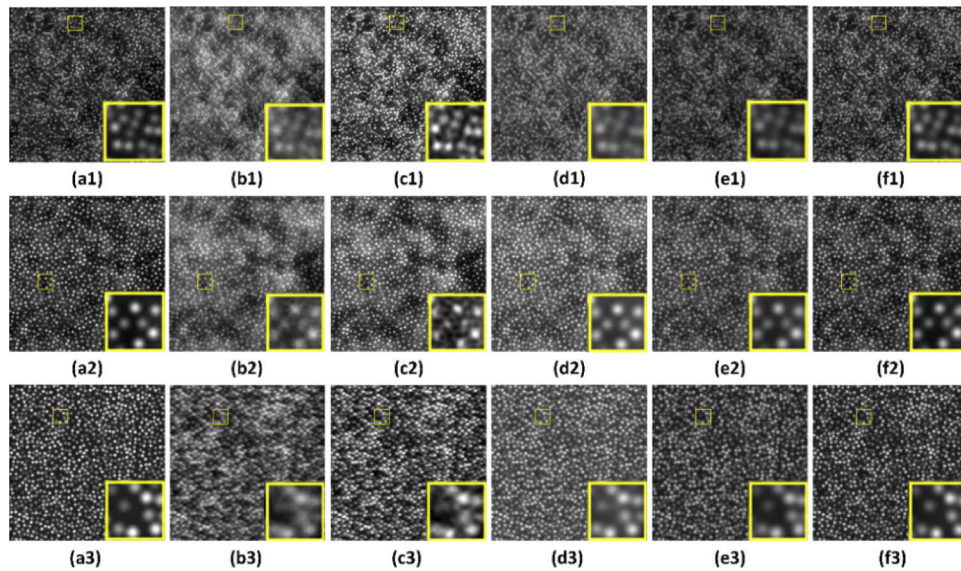
Since it is important for automated identification of photoreceptor cells for retinal images, the cone numbers are obtained by using the learning-based photoreceptor detection method proposed in [39]. The error rate of cone counting is computed as follows:

$$Error = \frac{|coneNo - truthNo|}{truthNo} \times 100\%, \quad (13)$$

where  $coneNo$  and  $truthNo$  is the cone numbers obtained from each deblurred image and the ground truth using the mentioned algorithm, respectively.

To demonstrate the usefulness of the proposed method, we compared the results of the proposed method with those of three state-of-the-art methods. One of the methods is the augmented Lagrangian method (ALM), which is a blind deconvolution method and frequently used as a comparison method in the AO retinal image deblurring literatures [10,12]. It is implemented from the publicly available code provided by the authors. The other methods are two natural scene image deblurring methods, DeblurGAN [18] and SRNdeblur [23]. These two models were both trained and tested on our synthetic retinal image dataset.

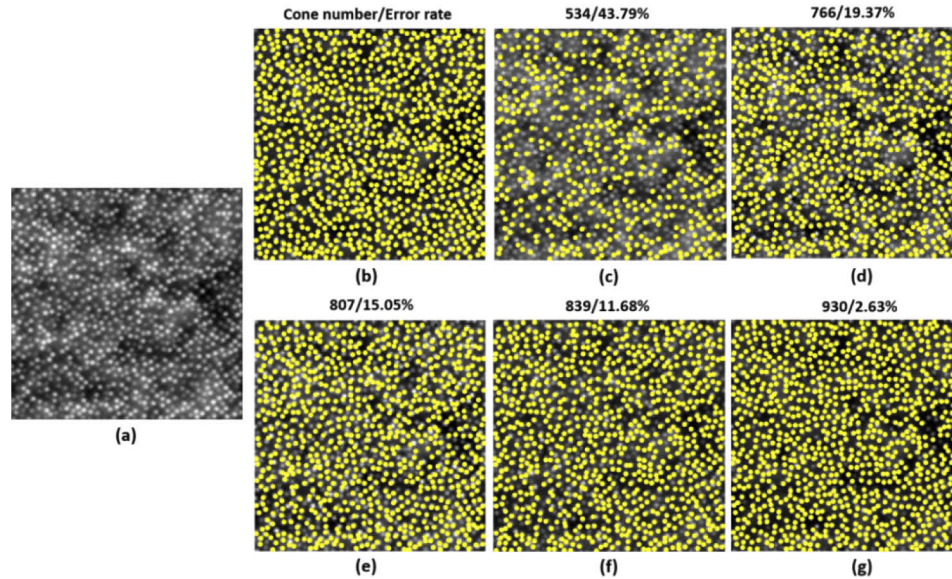
Figure 2 shows three typical examples of the original ideal images, corresponding blurry images (imitating at different eccentricities and noises), and the images restored by the three state-of-the-art methods and the proposed method. For better comparison, we zoom in on the local regions with yellow boxes. As illustrated in Figs. 2(c1)-(c3), the contrast of the images restored by the ALM has been significantly enhanced; however, the noises of the images have also been enhanced, and the restored images have obvious connection artifacts. Furthermore, this method performs unsatisfactorily when the images have large residual wavefront aberrations [Fig. 2(c3)]. As described in Figs. 2(d1)-(d3) and 2(e1)-(e3), the noises of the images restored by DeblurGAN and SRNdeblur were suppressed and the morphology of photoreceptor cells was restored; however, the images restored by these two methods were blurred to different extents, which make the photoreceptor cells difficult to be recognized. As shown in Figs. 2(f1)-(f3), the images restored by the proposed method have much sharper quality and more apparent photoreceptor cells, even when with large residual wavefront aberrations and noises.



**Fig. 2.** Deblurring results of synthetic AO retinal images imitating at different eccentricities and noises. (a1)-(a3) Ground truth; (b1) blurry images (0.3 mm eccentricity from the foveal center, and the Gaussian noise with a standard deviation of 0.03); (b2) blurry images (0.5 mm eccentricity from the foveal center, and the Gaussian noise with a standard deviation of 0.02); (b3) blurry images (1.0 mm eccentricity from the foveal center, and the Gaussian noise with a standard deviation of 0.05). Restored images from (c1), (c2), (c3) the ALM; (d1), (d2), (d3) the DeblurGAN method; (e1), (e2), (e3) the SRNdeblur method; and (f1), (f2), (f3) the proposed method. (Zooming-in the figure will provide a better look at the restoration quality).

Figures 3(a) and 4(a) present the blurry AO retinal image with small residual wavefront aberration and Gaussian noise with a standard deviation of 0.01, as well as the image blurred

by large residual wavefront aberrations and Gaussian noise with a standard deviation of 0.05. Figures 3(b)-(g) and 4(b)-(g) show the cone detection results of the ground truth and blurry image, and the cone detection results of the corresponding images restored by ALM, DeblurGAN, and SRNdeblur methods. It can be seen that the cone detection error rate of Fig. 4(d) is nearly two times higher than that of Fig. 3(d); this is because the ALM performed not well on the blurry image with large residual wavefront aberrations and noises. As shown in Figs. 3(e)-(g) and 4(e)-(g), the images deblurred by the proposed method have the lowest error rate; this indicates that the proposed method can generate images with sharper quality and that are more conducive to cone detection.

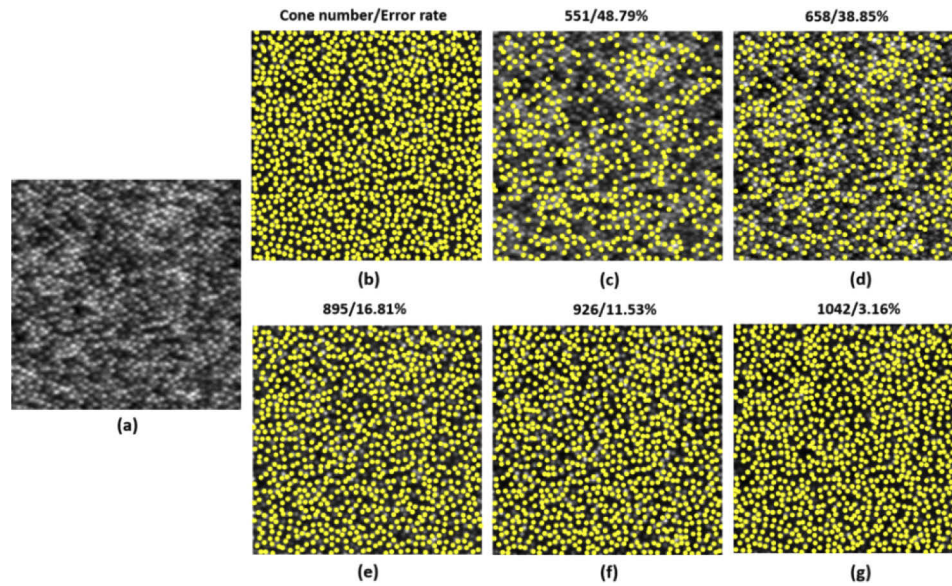


**Fig. 3.** Results of cone detection on synthetic retinal images. (a) Original blurry image (with small residual wavefront aberrations and Gaussian noise with a standard deviation of 0.01). Cone detection on (b) ground truth (1.5 mm eccentricity from the foveal center); (c) blurry image; image restored by (d) the ALM, (e) the DeblurGAN method, (f) the SRNdeblur method, and (g) the proposed method.

For quantitative comparison, the blurry images in the test set as well as the corresponding images restored by ALM, DeblurGAN, SRNdeblur, and the proposed method were evaluated with the metrics of PSNR, SSIM, perceptual distance, and error rate of cone counting. As seen in Table 1, the proposed method significantly outperforms the existing state-of-the-art image deblurring methods. The mean PSNR of the proposed method is 9.16 dB, 7.27dB, and 4.16dB higher than that of ALM, DeblurGAN, and SRNdeblur, respectively, and the mean SSIM is 0.24, 0.10 and 0.05 higher, respectively. Moreover, the mean perceptual distance of the proposed method is 0.0977, 0.0228, and 0.0156 lower than that of ALM, DeblurGAN, and SRNdeblur, respectively, and the mean error rate of cone counting is 26.69%, 12.89%, and 8.24% lower, respectively. Furthermore, we also compared the proposed method with the state-of-the-art AO deblurring method [10] which is called “AOdeblurCNN” here. As illustrated in Table 1, the mean PSNR of the proposed method surpassed that of the AOdeblurCNN 2.35 dB, and the mean SSIM surpassed 0.13. The mean error rate of cone counting of the proposed method is 2.76% lower than that of the AOdeblurCNN.

The above analysis shows that the proposed method can significantly improve the SNR and visual effect of images, which can facilitate better distinction of photoreceptor cells.





**Fig. 4.** Results of cone detection on synthetic retinal images. (a) Original blurry image (with large residual wavefront aberrations and Gaussian noise with a standard deviation of 0.05). Cone detection on (b) ground truth (1.0 mm eccentricity from the foveal center); (c) blurry image; image restored by (d) the ALM, (e) the DeblurGAN method, (f) the SRNdeblur method, and (g) the proposed method.

**Table 1. Performance comparison with state-of-the-art methods tested on synthetic retinal images (\* Data are expressed as mean  $\pm$  standard deviation)**

|                            | PSNR (dB)                          | SSIM                              | perceptual distance                   | Error rate of cone counting (%)   |
|----------------------------|------------------------------------|-----------------------------------|---------------------------------------|-----------------------------------|
| Blurry image               | $16.32 \pm 1.97$                   | $0.43 \pm 0.17$                   | $0.1407 \pm 0.0213$                   | $45.73 \pm 9.73$                  |
| ALM                        | $17.67 \pm 1.73$                   | $0.72 \pm 0.15$                   | $0.1106 \pm 0.0193$                   | $29.08 \pm 4.82$                  |
| DeblurGAN [18]             | $19.56 \pm 2.46$                   | $0.86 \pm 0.05$                   | $0.0357 \pm 0.0098$                   | $15.26 \pm 2.17$                  |
| SRNdeblur [23]             | $22.67 \pm 2.82$                   | $0.91 \pm 0.03$                   | $0.0285 \pm 0.0056$                   | $10.61 \pm 1.86$                  |
| AODEblurCNN [10]           | $24.48 \pm 2.76$                   | $0.83 \pm 0.07$                   | $\times$                              | $5.13 \pm 2.32$                   |
| <b>the proposed method</b> | <b><math>26.83 \pm 2.62</math></b> | <b><math>0.96 \pm 0.02</math></b> | <b><math>0.0129 \pm 0.0032</math></b> | <b><math>2.37 \pm 0.80</math></b> |

\*Since perceptual distance is of a smaller magnitude, it retains four significant digits, and other metrics retain two.

To compare the stability of each method, the coefficient of variance (CV) of different metrics for different methods has been calculated (Table 2). The CV represents the ratio of the standard deviation to the mean, and it is a useful statistic for comparing the degree of variance from one data series to another. A smaller CV indicates the better stability of the method. As illustrated in Table 2, the proposed method has the lowest CV in PSNR and SSIM, this implies that the proposed method has better stability in improving image quality. However, the CV of the proposed method in perceptual distance and error rate of cone counting is not the lowest; it is because the means of perceptual distance and error rate of cone counting are very small, if the standard deviations of them are not small enough, the value of CV will be high. Moreover, the values of these metrics vary greatly with the different images, thus the CV of perceptual distance and error rate of cone counting presented here is acceptable.

**Table 2. Statistical comparison of the CV of different metrics for different methods (\* Confidence interval is 95%)**

|                            | CV_PSNR (%) | CV_SSIM (%) | CV_perceptual distance (%) | CV_Error rate of cone counting (%) |
|----------------------------|-------------|-------------|----------------------------|------------------------------------|
| Blurry image               | 12.07       | 39.53       | 15.14                      | 21.27                              |
| ALM                        | 9.79        | 20.83       | 17.45                      | 16.57                              |
| DeblurGAN [18]             | 12.58       | 5.81        | 27.45                      | 14.22                              |
| SRNdeblur [23]             | 12.44       | 3.29        | 19.65                      | 17.53                              |
| AOdeblurCNN [10]           | 11.27       | 8.43        | ×                          | 45.22                              |
| <b>the proposed method</b> | <b>9.77</b> | <b>2.08</b> | <b>24.81</b>               | <b>33.76</b>                       |

### 3.2. Results evaluation for real AO retinal images

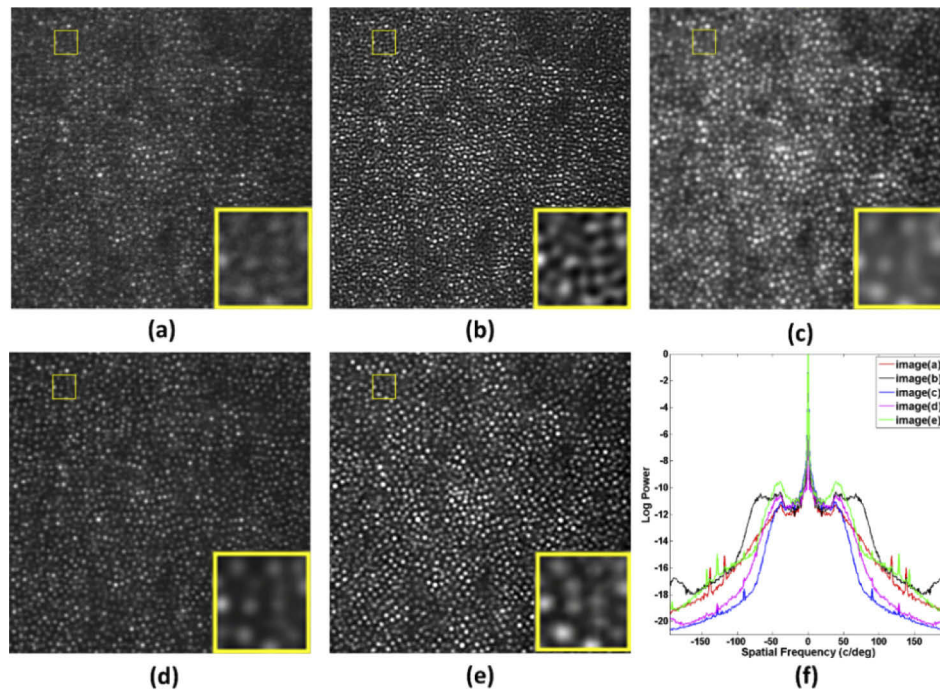
#### 3.2.1. Qualitative evaluation on real AO retinal images

Four hundred real AO retinal images were used to evaluate the effectiveness and generalization of the proposed method in image-quality improvement, of which 100 were captured from the AO system proposed in [40] with the field of view of 1.2 degrees and the retinal eccentricities of 0.8 mm and 1.5 mm from the foveal center, 300 were captured by the AO systems [41] in our laboratory with the field of view of 1.5 degrees and the retinal eccentricities ranged from 0.3 mm to 1.8 mm from the foveal center.

Figure 5(a) is an example of an original real retinal image with 0.8 mm eccentricity from the foveal center captured by the AO system proposed in [40]. As seen in Fig. 5(a), although the contours of the photoreceptor cells are visible, the serious noise makes the photoreceptor cells not easy to be identified. The noise of the images [Figs. 5(b)-(e)] restored by ALM, DeblurGAN, SRNdeblur, and the proposed method has been suppressed. However, according to the locally enlarged images of Figs. 5(b)-(e), the image restored by the ALM [Fig. 5(b)] has obvious connection artifacts, and the images restored by DeblurGAN and SRNdeblur methods [Figs. 5(c)-(d)] were blurred to different extents. As shown in Fig. 5(e), the image restored by the proposed method has sharper quality and the photoreceptor cells are effectively enhanced.

Figure 5(f) shows the power spectra of images in Figs. 5(a)-(e); the power spectra describe the amplitude of the spectral power in the images that is distributed across the spatial frequency. As shown in Fig. 5(f), the values of photoreceptor cell spatial frequencies of the images restored by the proposed method ranging from 50 to 75 c/deg are increased the most, which indicates that the visibility of photoreceptors are effectively improved. Since high frequencies probably corresponding to noise, as shown in Fig. 5(f), the power spectra of the images restored by DeblurGAN (blue curve) and SRNdeblur (pink curve) are rapidly decreased, which indicates that the noise of these restored images are effectively suppressed. However, the high-frequency power spectrum of the image restored by the proposed method (green curve) decreases slowly, and is even slightly higher than that of the original image (red curve). This is probably because the high-frequency represents the high-frequency details, which indicates that the details of the image restored by the proposed method have been enhanced; this can be proven by the inserts in Figs. 5(a)-(e).

Figure 6(a) is a typical example of an original real retinal image with 1.2 mm eccentricity from the foveal center captured from the AO system [41]. As shown in Fig. 6(a), since the imaging region is away from the foveal center, the photosensitivity of cells is decreased, and then the optical reflectivity of cells is decreased, which results in the residual aberrations increasing and the SNR of the AO retinal image decreasing. As illustrated in Figs. 6(b)-(e), the SNR of the images restored by ALM, DeblurGAN, SRNdeblur, and the proposed method are all improved. However, the image restored by ALM has connection artifacts, which seriously affect



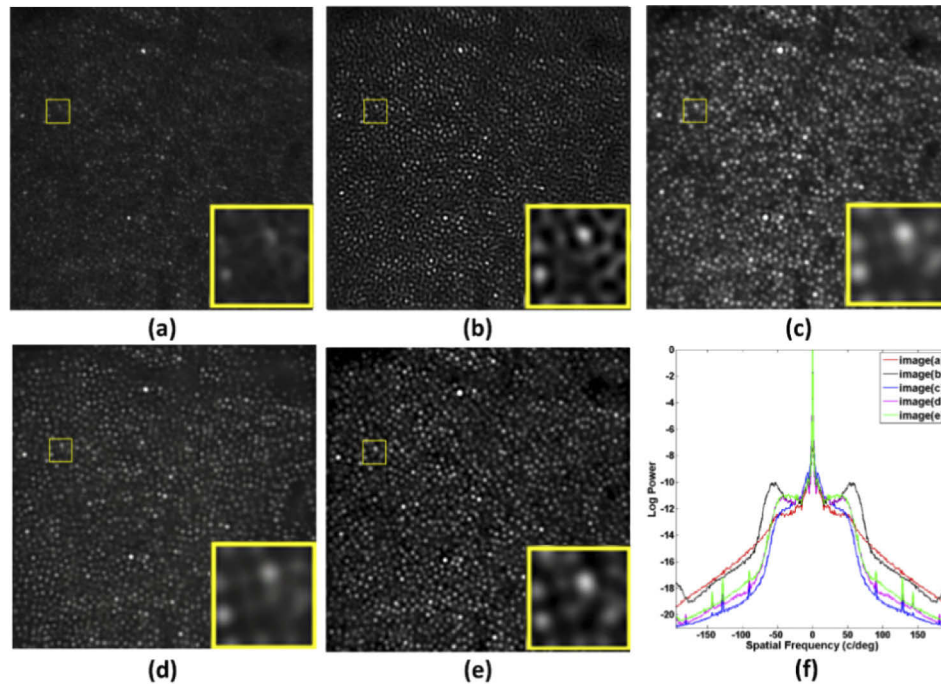
**Fig. 5.** Deblurring results of real retinal images captured by the AOSLO system [40]. (a) Original image (0.8 mm eccentricity from the foveal center). Restored images from (b) the ALM; (c) the DeblurGAN method; (d) the SRNdeblur method; and (e) the proposed method. (f) The corresponding image average power spectra. (Zooming-in the figure will provide a better look at the restoration quality).

the authenticity of the image. DeblurGAN and SRNdeblur methods can effectively smooth the images and enhance the brightness of the photoreceptor cells; however, it can be clearly seen in the inserts that the cells are blurred. Compared with the other three methods, the images restored by the proposed method have a better visual effect, and the residual wavefront aberrations of these images have been well corrected.

The power spectra of images in Figs. 6(a)-(e) are described in Fig. 6(f). As shown, the noises of the restored images are all suppressed. One should note that, compared with the images restored by DeblurGAN, SRNdeblur, and the proposed method (blue, pink, and green curve), the image restored by ALM (black curve) have the highest values of photoreceptor cell spatial frequencies when ranging from 50 to 75 c/deg. However, according to the details in the inset of Fig. 6(b), the image restored by ALM have obvious connection artifacts, which means the image restored by ALM method contains a lot of pseudo information. Moreover, the high-frequency power spectrum of the image restored by ALM (black curve) is just slightly lower than that of the original image (red curve), which means that the ALM performs unsatisfactorily in image denoising.

Figure 7(a) is the example of original real retinal image with 1.8mm eccentricity from the foveal center captured from the AO system [41]. The imaging regions of this image is further away from the foveal center, thus the image is affected by edge aberration and makes a lower image quality, which is more in need of image quality improvement. The contrast of the image restored by ALM [Fig. 7(b)] have been significantly enhanced, however, these images also with obvious connection artifacts. Furthermore, according to the power spectrum curve of Fig. 7(b) (Fig. 7(f), black curve), ALM performs not well in image denoising when comparing with other three methods. The power spectrum of Fig. 7(c) is shown as the blue curve of Fig. 7(f), and it





**Fig. 6.** Deblurring results of real retinal images captured by the AOSLO system [41]. (a) Original image (1.2 mm eccentricity from the foveal center). Restored images from (b) the ALM; (c) the DeblurGAN method; (d) the SRNdeblur method; and (e) the proposed method. (f) Corresponding image average power spectra. (Zooming-in the figure will provide a better look at the restoration quality).

can be seen, although DeblurGAN has the best performance in denoising, it cannot effectively enhance the low-frequency information, and thus the photoreceptors cannot be well detected. Comparing the power spectrum of the image restored by SRNdeblur with that restored by the proposed method, we can find that, these two methods both perform well in noise suppression and SNR improvement, however, according to the locally enlarged images and the overall visual effect of the images, the image restored by the proposed method has higher SNR and more visible photoreceptor cells.

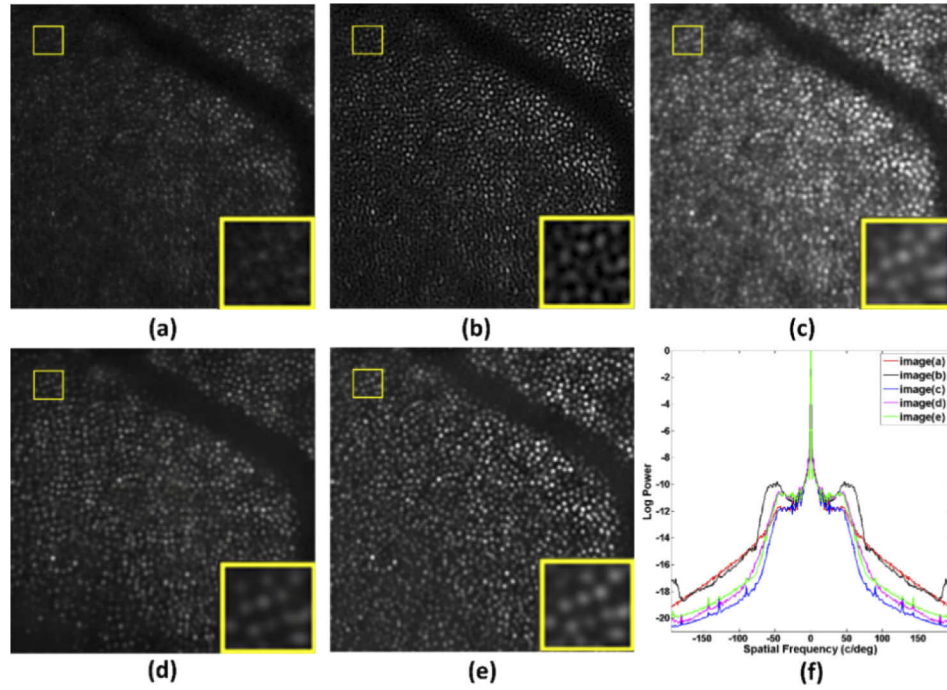
In summary, according to the visual effect and the power spectra of the restored images, we can conclude that compared with the images whose residual aberrations were well corrected [Fig. 5(a)], the proposed method performs more remarkably on the images seriously affected by residual aberrations [Fig. 6(a) and 7(a)].

### 3.2.2. Quantitative evaluation on real retinal images

For real retinal images, there was no ground truth image for the evaluation of the results. Thus, a no-reference image quality assessment (NR-IQA) algorithm was needed. In this work, blind image quality index (BIQI) [42] was used to quantitatively evaluate the method's performance.

BIQI is a commonly used method for quantitatively evaluating the quality of no-reference images. This method has a two-step framework. Given an image, the algorithm first estimates the presence of a set of distortions in the image; the set consists of JPEG, JPEG2000, white noise, Gaussian blur and fast fading. Then, the quality of the image along each of these distortions can





**Fig. 7.** Deblurring results of real retinal images captured by AOSLO system [41]. (a) Original image (1.8 mm eccentricity from the foveal center). Restored images from (b) the ALM; (c) the DeblurGAN method; (d) the SRNdeblur method; and (e) the proposed method. (f) The corresponding image average power spectra. (Zooming-in the figure will provide a better look at the restoration quality).

be evaluated, and the quality of image can be expressed as follows:

$$BIQI = \sum_{i=1}^5 p_i \cdot q_i \quad (14)$$

where  $p_i$  ( $i=1 \dots 5$ ) represents the probability of each distortion in the image, and  $q_i$  ( $i=1 \dots 5$ ) denotes the quality scores from each of the five quality assessment algorithms corresponding to the five distortions.

For an input image, a BIQI score ranging from 0 to 100 can be obtained using this method. A lower BIQI score indicates better image quality. The BIQI scores of the 400 real retinal images captured from different AO systems [40,41] were calculated. As shown in Table 3, the proposed method outperforms the other four methods. The BIQI score of the proposed method is  $36.74 \pm 4.72$ , which is significantly lower than that of the other four methods. This indicates that the images restored by the proposed method have better quality.

**Table 3. Performance comparison with state-of-the-art method tested on real retinal images (\* Data are expressed as mean  $\pm$  standard deviation)**

|            | Original real image | ALM              | DeblurGAN [18]   | SRNdeblur [23]   | AODEblurCNN [10] | the proposed method              |
|------------|---------------------|------------------|------------------|------------------|------------------|----------------------------------|
| BIQI score | 60.94 $\pm$ 8.09    | 55.86 $\pm$ 6.74 | 48.23 $\pm$ 5.12 | 46.18 $\pm$ 5.08 | 51.62 $\pm$ 4.02 | <b>36.74<math>\pm</math>4.72</b> |

As described and analyzed in section 3.2.1, we can find that the performance of the methods may vary by different systems and eccentricities. Here, the performance of the methods under different systems and eccentricities is quantitatively evaluated, respectively.

## (1) Quantitative evaluation of the performance under different systems

To avoid the influence of the real retinal images with different eccentricities, we selected 20 real images with an eccentricity of 0.8 mm from the images captured by the two systems [40,41], respectively. These images are restored by ALM, DeblurGAN, SRNdeblur, and the proposed method, and their statistical results are illustrated in Table 4. The BIQI score difference between original real images and restored images are shown in the parentheses. According to the statistical results in Table 4, we have the following findings. (i) The BIQI score of the images captured by system [40] is lower than that captured by system [41], this implies the image captured by system [40] has better quality. (ii) The three state-of-the-art methods performed better on the images captured by system [40], whereas the proposed method performs better on the images captured by system [41].

## (2) Quantitative evaluation of the performance under different eccentricities

To better evaluate the image restoration performance vary by eccentricities, here we only use the images captured by system [41]. As known from the distribution characteristics of photoreceptor cells [43–45], a smaller distance to the foveal center implies more photoreceptor cells. Since the retinal eccentricities of the images in the real dataset ranged from 0.3 mm to 1.8 mm from the foveal center, we select 20 images with eccentricity of 0.3 mm, 20 with eccentricity of 1.2 mm, and 20 with eccentricity of 1.8 mm to evaluate the performance vary by eccentricities. The statistical results are illustrated in Table 5, and the BIQI score difference between original real images and restored images are shown in the parentheses. According to the statistical results, we have the following findings. (i) The image with smaller eccentricity has higher image quality. (ii) The four methods perform more remarkable on the image with large eccentricity, however, the proposed method has the best performance.

**Table 4. Image restoration results comparison for the real retinal images captured from different systems (\* Data are expressed as mean  $\pm$  standard deviation)**

|                        | Original real image | ALM ( $\Delta$ difference)            | DeblurGAN [18] ( $\Delta$ difference)  | SRNdeblur [23] ( $\Delta$ difference)  | the proposed method ( $\Delta$ difference) |
|------------------------|---------------------|---------------------------------------|--|--|--|
| Image from system [40] | 56.39 $\pm$ 2.24    | 52.25 $\pm$ 2.13<br>(4.14 $\pm$ 0.11) | 46.36 $\pm$ 2.10<br>(10.03 $\pm$ 0.14) | 43.11 $\pm$ 1.97<br>(13.28 $\pm$ 0.27) | 34.03 $\pm$ 1.86<br>(22.36 $\pm$ 0.38)     |
| Image from system [41] | 57.76 $\pm$ 2.34    | 54.35 $\pm$ 2.17<br>(3.41 $\pm$ 0.17) | 48.21 $\pm$ 1.97<br>(9.55 $\pm$ 0.37)  | 44.75 $\pm$ 2.03<br>(13.00 $\pm$ 0.31) | 34.56 $\pm$ 1.92<br>(23.20 $\pm$ 0.42)     |

**Table 5. Image restoration results comparison for the real retinal images with different eccentricities (\* Data are expressed as mean  $\pm$  standard deviation)**

|                                | Original real image | ALM ( $\Delta$ difference)            | DeblurGAN [18] ( $\Delta$ difference)  | SRNdeblur [23] ( $\Delta$ difference)  | the proposed method ( $\Delta$ difference) |
|--------------------------------|---------------------|---------------------------------------|--|--|--|
| Image with 0.3 mm eccentricity | 55.27 $\pm$ 2.42    | 51.29 $\pm$ 2.17<br>(3.98 $\pm$ 0.25) | 44.85 $\pm$ 1.74<br>(10.42 $\pm$ 0.68) | 42.17 $\pm$ 1.07<br>(13.1 $\pm$ 1.35)  | 32.72 $\pm$ 0.70<br>(22.55 $\pm$ 1.72)     |
| Image with 1.2 mm eccentricity | 63.11 $\pm$ 2.47    | 57.55 $\pm$ 1.86<br>(4.56 $\pm$ 0.61) | 51.83 $\pm$ 1.62<br>(11.28 $\pm$ 0.85) | 47.25 $\pm$ 1.34<br>(15.86 $\pm$ 1.13) | 37.79 $\pm$ 0.72<br>(25.32 $\pm$ 1.75)     |
| Image with 1.8 mm eccentricity | 66.96 $\pm$ 2.07    | 60.87 $\pm$ 1.73<br>(5.59 $\pm$ 0.84) | 54.85 $\pm$ 1.50<br>(11.61 $\pm$ 1.07) | 49.95 $\pm$ 1.31<br>(16.51 $\pm$ 1.26) | 40.39 $\pm$ 0.71<br>(26.57 $\pm$ 1.36)     |

All above the statistical comparison results demonstrate that the proposed method can effectively improve the image quality, especially the image with low quality, which agree with the qualitative analyses shown in section 3.2.1 to some extent.

### 3.3. Time complexity

Floating-point operations (FLOPs) were used to measure the time complexity of the model. The formulas of FLOPs are described as follows:

$$\text{Time} \sim O\left(\sum_{l=1}^D M_l^2 \cdot K_l^2 \cdot C_{lin} \cdot C_{lout}\right) \quad (15)$$

$$M = \frac{(N - K + 2 \times P)}{S} + 1 \quad (16)$$

In formula (15),  $D$  is the number of convolutional layers;  $M_l$  is the output feature map of the  $l$ th layer;  $K_l$  is the convolution kernel size of the  $l$ th layer;  $C_{lin}$  is the input channel of the  $l$ th layer;  $C_{lout}$  is the output channel of the  $l$ th layer. In formula (16),  $N$  is the size of input image;  $K$  is the convolution kernel size; Padding is expressed as  $P$ , which equals to zero if padding is “valid”, otherwise  $M$  directly equals to  $N$  when padding is “same”.  $S$  represents the step length of filter in both vertical and horizontal directions in the original image.

Since AOdeblurCNN was the first fully deep-learning-based method applied to AO image deblurring task, here we calculate and compare the time complexity of the proposed method with that of AOdeblurCNN. The configuration details of the proposed model and the AOdeblurCNN are shown in Tables 6–8 of the appendix part (A.1 and A.2). The FLOPs of the proposed generator and discriminator is  $1.78 \times 10^6 N^2$  and  $1.72 \times 10^4 N^2$ , and the total time complexity of the proposed model is  $1.797 \times 10^6 N^2$ , where  $N$  is the size of the input image. Meanwhile, the FLOPs of AOdeblurCNN is  $2.12 \times 10^5 N^2$ . Although the time complexity of the proposed method is larger than the state-of-the-art AO deblurring method, the time it takes to restore an image with  $256 \times 256$  pixels on an ordinary CPU is comparable, both methods taking about 1 second. Moreover, the proposed method can generate better results, and the training of our model takes place offline, so the time complexity of the proposed model is acceptable.

## 4. Discussion and conclusion

Adaptive optics (AO) retinal imaging changes the way we look at the retina in an intact living eye. AO can provide optical resolutions of  $2 \mu\text{m}$  or less in the human eye, and this resolution is sufficient to make measurements of cellular and sub-cellular details of the normal retina structure. However, when affected by residual aberration, intraocular scatter, speckle noise, and other factors, the quality of AOSLO images is degraded and the photoreceptor cells are obscured. In recent years, many efforts have been made to improve the quality of retinal images, and with the development of deep learning, the deep-learning-based methods have also been applied to the improvement of AO retinal images. In [10], the authors adopted the CNN to restore AO retinal images, which was the first fully deep-learning-based method applied to this problem. In this study, we proposed a fully deep-learning-based method based on the conditional GAN to directly learn the mapping relation between blurry and restored images. In the training of the proposed model, the pixel-space loss and feature-space loss were introduced to help restore images of high perceptual quality. Furthermore, the PatchGAN-based discriminator was designed to help reconstruct image details and improve training efficiency.

To demonstrate the effectiveness of the proposed model, we compared the proposed method with the state-of-the-art natural scene image deblurring methods [18,23], both qualitatively and quantitatively on synthetic retinal images, as well as on real retinal images captured from different systems. Besides, we also compared the proposed method with the state-of-the-art AO retinal image deblurring method [10] on the quantitative evaluation metrics and time complexity. Since there is no publicly available code of AOdeblurCNN, its quantitative results were all from the description of their article; this result may be somewhat different from the result of their model training and testing directly on our dataset, but the way we used for generating synthetic retinal





**Table 7. Configuration detail of the convolutional layers in the proposed discriminator network**

| Layer        | Input Feature Size   | Output Feature Size (M) | Kernel Size (K) | Input Channel (Cin) | Output Channel (Cout) | Stride (S) | Padding (P) |
|--------------|----------------------|-------------------------|-----------------|---------------------|-----------------------|------------|-------------|
| Conv1        | 256                  | 128                     | 4               | 3                   | 64                    | 2          | 1           |
| Conv2        | 128                  | 64                      | 4               | 64                  | 128                   | 2          | 1           |
| Conv3        | 64                   | 32                      | 4               | 64                  | 256                   | 2          | 1           |
| Conv4        | 32                   | 16                      | 4               | 128                 | 512                   | 2          | SAME        |
| FC           | 16                   | 1                       | -               | -                   | -                     | -          | -           |
| <b>FLOPs</b> | <b>1,124,073,472</b> |                         |                 |                     |                       |            |             |

*A.2. Configuration details of the state-of-the-art AO deblurring model***Table 8. Configuration detail of the convolutional layers in the state-of-the-art AO deblurring model [10]**

| Layer        | Input Feature Size    | Output Feature Size (M) | Kernel Size (K) | Input Channel (Cin) | Output Channel (Cout) | Stride (S) | Padding |
|--------------|-----------------------|-------------------------|-----------------|---------------------|-----------------------|------------|---------|
| Conv1        | 256                   | 246                     | 11              | 3                   | 64                    | 1          | VALID   |
| Conv2        | 246                   | 244                     | 3               | 64                  | 64                    | 1          | VALID   |
| Conv3        | 244                   | 242                     | 3               | 64                  | 64                    | 1          | VALID   |
| Conv4        | 242                   | 238                     | 3               | 64                  | 64                    | 1          | VALID   |
| Conv5        | 238                   | 234                     | 5               | 64                  | 64                    | 1          | VALID   |
| <b>FLOPs</b> | <b>13,907,378,176</b> |                         |                 |                     |                       |            |         |

**Funding**

National Basic Research Program of China (2016YFE0107700, 2017YFC0108201); Jiangsu Provincial Key Research and Development Program (BE2018667, BE2019682); National Natural Science Foundation of China (61605210, 61675226); Youth Innovation Promotion Association of the Chinese Academy of Sciences (2019320); Strategic Priority Research Program of Chinese Academy of Sciences (XDB32000000).

**Disclosures**

The authors declare that there are no conflicts of interest related to this article.

**References**

1. J. Liang, D. R. Williams, and D. T. Miller, "Supernormal vision and high-resolution retinal imaging through adaptive optics," *J. Opt. Soc. Am. A* **14**(11), 2884–2892 (1997).
2. F. Vargasmartin, P. M. Prieto, and P. Artal, "Correction of the aberrations in the human eye with a liquid-crystal spatial light modulator: limits to performance," *J. Opt. Soc. Am. A* **15**(9), 2552–2562 (1998).
3. A. Roorda, F. Romeroborja, W. J. Donnelly, H. M. Queener, T. J. Hebert, and M. C. W. Campbell, "Adaptive optics scanning laser ophthalmoscopy," *Opt. Express* **10**(9), 405–412 (2002).
4. S. A. Burns, S. Marcos, A. E. Elsner, and S. Bara, "Contrast improvement of confocal retinal imaging by use of phase-correcting plates," *Opt. Lett.* **27**(6), 400–402 (2002).
5. J. Lu, H. Li, L. Wei, G. Shi, and Y. Zhang, "Retina imaging in vivo with the adaptive optics confocal scanning laser ophthalmoscope," in *IEEE International Conference on Photonics* (2009), 7519.
6. J. Arines, "Partially compensated deconvolution from wavefront sensing images of the eye fundus," *Opt. Commun.* **284**(6), 1548–1552 (2011).
7. V. Nourrit, B. Vohnsen, and P. Artal, "Blind deconvolution for high-resolution confocal scanning laser ophthalmoscopy," *J. Opt.* **7**(10), 585–592 (2005).

8. J. C. Christou, A. Roorda, and D. R. Williams, "Deconvolution of adaptive optics retinal images," *J. Opt. Soc. Am. A* **21**(8), 1393–1401 (2004).
9. H. Li, J. Lu, G. Shi, and Y. Zhang, "Real-time blind deconvolution of retinal images in adaptive optics scanning laser ophthalmoscopy," *Opt. Commun.* **284**(13), 3258–3263 (2011).
10. X. Fei, J. Zhao, H. Zhao, D. Yun, and Y. Zhang, "Deblurring adaptive optics retinal images using deep convolutional neural networks," *Biomed. Opt. Express* **8**(12), 5675–5687 (2017).
11. C. Rao, Y. Tian, and H. Bao, *AO-Based High Resolution Image Post-Process* (InTech, 2012).
12. A. Lazareva, M. Asad, and G. G. Slabaugh, "Learning to Deblur Adaptive Optics Retinal Images," in *International Conference on Image Analysis and Recognition* (2017), pp. 497–506.
13. M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 1–14.
14. D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 2536–2544.
15. R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in *European Conference on Computer Vision (ECCV)* (Springer, 2016), pp. 649–666.
16. W. Li, Y. He, W. Kong, F. Gao, J. Wang, and G. Shi, "Enhancement of Retinal Image From Line-Scanning Ophthalmoscope Using Generative Adversarial Networks," *IEEE Access* **7**, 99830–99841 (2019).
17. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems* (2014), pp. 2672–2680.
18. O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 8183–8192.
19. X. Yu, Y. Qu, and M. Hong, "Underwater-GAN: Underwater Image Restoration via Conditional Generative Adversarial Network," in *International Conference on Pattern Recognition (ICPR)* (2018), pp. 66–75.
20. S. Osindero, "Conditional Generative Adversarial Nets," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014), pp. 1–7.
21. P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 1–17.
22. C. Li and M. Wand, "Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks," in *European Conference on Computer Vision (ECCV)* (2016), pp. 702–716.
23. X. Tao, H. Gao, Y. Wang, X. Shen, J. Wang, and J. Jia, "Scale-recurrent Network for Deep Image Deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 1–9.
24. L. Mariotti and N. Devaney, "Performance analysis of cone detection algorithms," *J. Opt. Soc. Am. A* **32**(4), 497–506 (2015).
25. L. N. Thibos, A. Bradley, and X. Hong, "A statistical model of the aberration structure of normal, well-corrected eyes," *Oph. Phys. Optics* **22**(5), 427–433 (2002).
26. D. T. Miller and A. Roorda, "Adaptive Optics in Retinal Microscopy and Vision," in *Handbook of Optics* (McGraw-Hill, 2009).
27. D. X. Hammer, R. D. Ferguson, C. E. Bigelow, N. Iftimia, T. E. Ustun, and S. A. Burns, "Adaptive optics scanning laser ophthalmoscope for stabilized retinal imaging," *Opt. Express* **14**(8), 3354–3367 (2006).
28. R. F. Cooper, C. S. Langlo, A. Dubra, and J. Carroll, "Automatic detection of modal spacing (Yellott's ring) in adaptive optics scanning light ophthalmoscope images," *Ophthalmic Physiol. Opt.* **33**(4), 540–549 (2013).
29. S. A. Burns, R. Tumbar, A. E. Elsner, D. Ferguson, and D. X. Hammer, "Large Field of View, Modular, Stabilized, Adaptive-Optics-Based Scanning Laser Ophthalmoscope," *J. Opt. Soc. Am. A* **24**(5), 1313–1326 (2007).
30. Y. He, G. Deng, L. Wei, X. Li, J. Yang, G. Shi, and Y. Zhang, "Design of a Compact, Bimorph Deformable Mirror-Based Adaptive Optics Scanning Laser Ophthalmoscope," *Adv. Exp. Med. Biol.* **923**, 375–383 (2016).
31. M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 1–32.
32. Z. Chen and Y. Tong, "Face super-resolution through wasserstein gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 1–8.
33. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. P. Aitken, A. Tejani, J. Totz, and Z. Wang, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 105–114.
34. S. Nah, T. H. Kim, and K. M. Lee, "Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 257–265.
35. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014), pp. 1–14.
36. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of International Conference for Learning Representations (ICLR)* (2015), pp. 1–15.
37. R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 586–595.

38. A. Lucas, S. L. Tapia, R. Molina, and A. K. Katsaggelos, "Generative Adversarial Networks and Perceptual Losses for Video Super-Resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 51–55.
39. B. Davidson, A. Kalitzeos, J. Carroll, A. Dubra, S. Ourselin, M. Michaelides, and C. Bergeles, "Automatic Cone Photoreceptor Localisation in Healthy and Stargardt Afflicted Retinas Using Deep Learning," *Sci. Rep.* **8**(1), 7911 (2018).
40. J. Lu, B. Gu, X. Wang, and Y. Zhang, "High speed adaptive optics ophthalmoscopy with an anamorphic point spread function," *Opt. Express* **26**(11), 14356–14374 (2018).
41. Y. Wang, Y. He, L. Wei, X. Li, J. Yang, H. Zhou, and Y. Zhang, "Bimorph deformable mirror based adaptive optics scanning laser ophthalmoscope for retina imaging in vivo," *Chin. Opt. Lett.* **15**(12), 121102 (2017).
42. A. K. Moorthy and A. C. Bovik, "A Two-Step Framework for Constructing Blind Image Quality Indices," *IEEE Signal Process. Lett.* **17**(5), 513–516 (2010).
43. H. Song, C. T. Y. Ping, Z. Zhangyi, A. E. Elsner, and S. A. Burns, "Variation of Cone Photoreceptor Packing Density with Retinal Eccentricity and Age," *Invest. Ophthalmol. Visual Sci.* **52**(10), 7376–7384 (2011).
44. M. N. Muthiah, C. Gias, F. K. Chen, J. Zhong, Z. McClelland, F. B. Sallo, T. Peto, P. J. Coffey, and L. da Cruz, "Cone photoreceptor definition on adaptive optics retinal imaging," *Br. J. Ophthalmol.* **98**(8), 1073–1079 (2014).
45. B. A. Wandell, *Foundations of Vision* (Sinauer Associates Inc, 1995), Chap.3.